

# ER390 Project Proposal

## Visualizing HAT's GOE Monitoring Data

Wendy Anthony wanthonny@uvic.ca

2025/09/09

## Contents

<b>The Idea Behind This Proposal</b>	<b>2</b>
<b>Background Research</b>	<b>2</b>
<b>Project Partner</b>	<b>2</b>
<b>Project Goals and Overall Purpose</b>	<b>2</b>
<b>Project Objectives</b>	<b>2</b>
<b>Project Methods</b>	<b>3</b>
<b>Project Deliverables</b>	<b>3</b>
<b>Workflow Graph</b>	<b>4</b>
<b>Proposed Timeline</b>	<b>5</b>
<b>Data Cleaning and Transformation Example Results</b>	<b>6</b>
Data download and formatting . . . . .	6
Load Site Data . . . . .	6
Data Cleaning and Transformation . . . . .	6
<b>Data Visualization Example Results</b>	<b>7</b>
Data Table . . . . .	7
Map of Site Locations . . . . .	8
Subregion Plots . . . . .	9
Site Plots . . . . .	10

## The Idea Behind This Proposal

After hearing Nancy Shackelford's suggestion to create a data visualization tool for HAT Garry Oak Ecosystems (GOE) Monitoring as a suitable ER390 Final Project, I began thinking about several creative solutions. Using example data from tables in the ECCC Monitoring report that influenced HAT's project design (Malloff & Shackelford, 2024), I began to develop R code, modifiable to incorporate HAT data when available. This case study may also model methods enabling comparison of other Garry Oak Ecosystem Monitoring Projects.

## Background Research

Comparing monitoring results from multiple restoration projects requires consistent measurements of vegetation composition (e.g. native and exotic species cover), which may help identify responses to stressors affecting already compromised and rare GOE. Representing data in a visual manner can help make complex data more understandable than a tables of numbers, and may reveal patterns or trends resulting from different restoration management actions (Malloff & Shackelford, 2024). Data visualization can represent abstract information visually, allow for data exploration and analysis, help model restoration management alternatives, create "data-based story-telling", and communicate potential insights to help make effective restoration management decisions; RMarkdown-generated reports can help create reproducible research (McKay, 2020).

° Malloff, J., & Shackelford, N. (2024). *Feeling the Pulse: Monitoring methods and initial outcomes in oak meadow ecosystems*. Restoration Futures Lab at the University of Victoria.

° McKay, S. K. (2020). *Data visualization for ecological analysis and restoration*. US Army Corp of Engineers:Ecosystem Management and Restoration Research Program (EMRRP) Webinar Series. <https://cw-environment.erdc.dren.mil/webinars/20Feb5-EcoMod-DataViz.pdf>

## Project Partner

I spoke to Vanessa Brownlee, Habitat Restoration Coordinator at Habitat Acquisition Trust (HAT), during their 2025-08-27 Social Gathering, who expressed interest in my proposed project idea, wanted me to send her more information about what I had in mind, and seemed willing to consider me starting in Fall 2025, with the understanding that interactions will be limited and data will not be available until later in October. I also later spoke briefly to Board Chair Andy MacKinnon, who seemed keen to know of my interest in doing my RNS Final Project with HAT and looks forward to learning more. I'm prepared to be guided by HAT's goals, preferences, timing, and data protocols, with Nancy Shackelford's RNS guidance and support.

## Project Goals and Overall Purpose

I propose to create a customized coding and reporting tool for non-coders, with tutorial instructions, using open-source R packages to format and interactively visualize data collected at multiple sites for HAT Garry Oak Ecosystem Monitoring (GOEM) Program. Based on questions HAT wants the tool to help answer about the data, it will provide an interactive way to communicate information about the data and help determine any patterns through analysis. This project proposal was created using R code in RMarkdown, to showcase some visualizations tools and reports that could be created, after adapting code to reflect HAT data. This may also be a method model to compare multiple GOEM Projects using the same measurement parameters.

## Project Objectives

- (See *WorkFlowchart, Figure 1 & Gantt Chart Proposed Timeline, Figure 2*)
- **Analyze and transform collected data** to understand data, structure, ensure consistent naming protocols, format data for code use, create code to clean raw data, calculate metrics.
- **Write data visualization code** for HAT data using RStudio, RMarkdown (Rmd), and RShiny, using reproducible code to create appropriate visualizations to enable a better understanding of the collected

data; for use by non-coders; download, format, save data and results; produce customized reports; create custom R functions to read, analyze, and plot data, as basis for proposed HAT `goemR` package.

- **Create interactive data visualization tools** Scaleable, to allow for more sites to be added; multi-year data; different measuring and sensitivity criteria (e.g. First Nations GOE sites); customize for consistent styles using HAT colours, logos, and plot themes.
- **Create interactive, browser-based Shiny App** User-chosen, custom-made data visualization plots, maps, and statistical analysis; select location(s) and variables; upload data; download results. Interactive choices will react to other variable selections to make custom-made visualizations and analysis.
- **Develop methods to allow for comparison of other GOE Monitoring Projects Data** Using the same measurement parameters and metrics (Malloff & Shackelford, 2024).
- **Analyze results patterns** of individual site differences with variables, metrics and indices.

## Project Methods

- (See *WorkFlowchart, Figure 1 & Gantt Chart Proposed Timeline, Figure 2*)
- Continue research for Shiny reactive code to make interactive data choices (e.g. drop-down lists), respond to choices, and filter results.
- Code in Rmd test files until code runs correctly, before copying code to working Rmd file and Shiny App. Develop functional code chunks one-at-a-time before starting next parts; troubleshoot small sections to get a fully working product before adding more functionality; Code is commented for communication and reproducibility; Create reusable code functions as basis for R Package
- Create random ECCC 2nd year test data for producing multi-year, multi-project comparisons
- Sign any HAT data confidentiality agreements as required
- Consult with HAT restoration monitor team to understand collected data, formats (forms, spreadsheets, csv), variables measured, indices to be calculated, questions to explore with data to analyze results
- Analyze, clean, and transform data, check for inconsistencies; calculate indices (See *Data Code*)
- Create data visualization plot charts that best reflect nature of variables being measured
- Create HAT-customized colour palettes, report and plot themes
- Create portable Project folder, with subfolders: R scripts, raw data, processed data, images, reports, meta data (units, data descriptions, etc); create naming protocols; version-saving with GitHub
- Document methods, successes, and failures in the process of the project for final paper
- Continued research for using reproducible R code, and of effective use of data visualization in ecological restoration monitoring, for inclusion in final report

## Project Deliverables

- (See *WorkFlowchart, Figure 1 & Proposed Timeline, Figure 2*); RStudio-created Deliverables
- **RMarkdown Data Processing Document Tool:** Data download, cleaning, transformation for visualization code used in Shiny App; reports: HTML, pdf, Word, with plot output, or code snippets
- **Customizations:** Create HAT-customized colour palette for Plot themes, App, Reports; add logo
- **Create Customized Functions for HAT** to form core of a HAT-specific R Package library
- **Create Report Templates:** Customized for individual site, total project reports, or at HAT's request
- **Data Visualizations:** Including data tables, plots, interactive maps; for subgroups, individual sites, site comparison, or for confidential use by First Nations (See *Example Visualization*)
- **Shiny App Interactive Data Viz Tool:** Data processed in Rmd; choose variables to visualize tables, plot, maps, stats; run local for confidentiality (See *Shiny App Sampler*)
- **Tutorials for using Data visualization Tools:** customized, easy to understand
- **Best practises for data files:** To ensure naming consistency in files, variables, folder structure, etc.
- **Presentation:** Present completed project to community partner HAT; including slides, live demonstrations of Data Visualization Tools, and written report after approval for publishing.
- **Written Report:** Due 2026-08-15, document project, using reproducible R code, and visualizing ecological restoration monitoring data; publish in online journal Ecorestoration: RNS Technical Series

# Workflow Graph

- (See **Objectives**, **Methods**, and **Deliverables**)
- The workflow shows how the raw data, from HAT monitoring measurements, would be transformed and included in code to produce a variety of deliverable data visualization products.
- This graph is made with R code from DiagrammeR, using a colour-blind friendly colour palette.

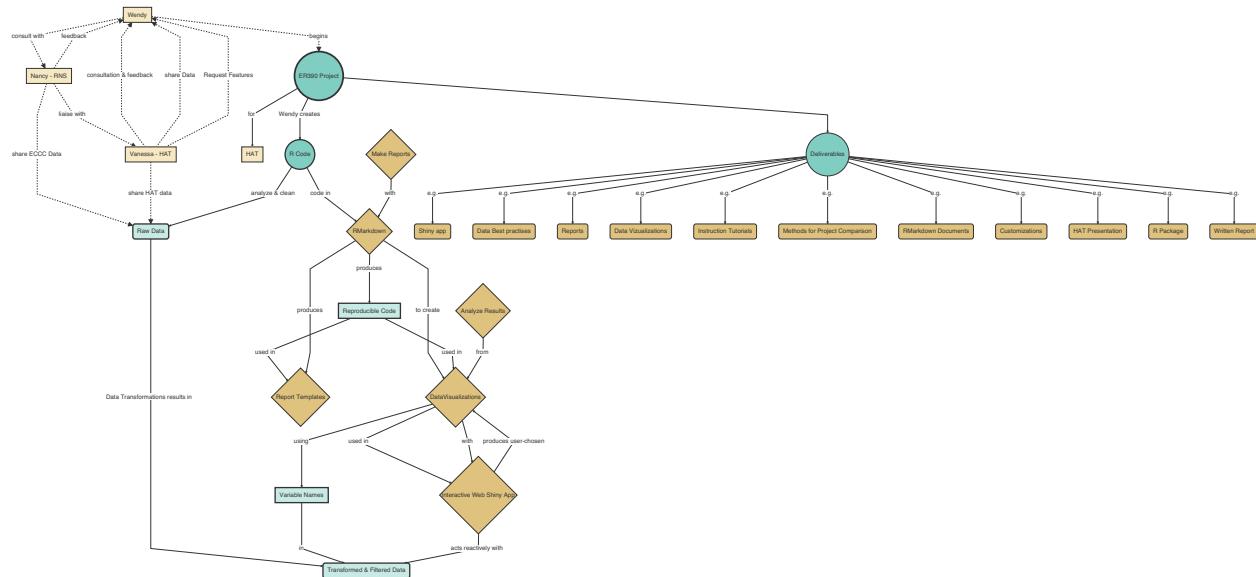


Figure 1: Project method workflow diagram for ER390 data visualization deliverables for HAT

# Proposed Timeline

- red coloured items are considered critical
- grey items are completed
- Note: The Gantt Chart is coded with DiagrammeR using imported csv file of data, but has no colour or size control
- (*See Objectives, Methods, and Deliverables*)

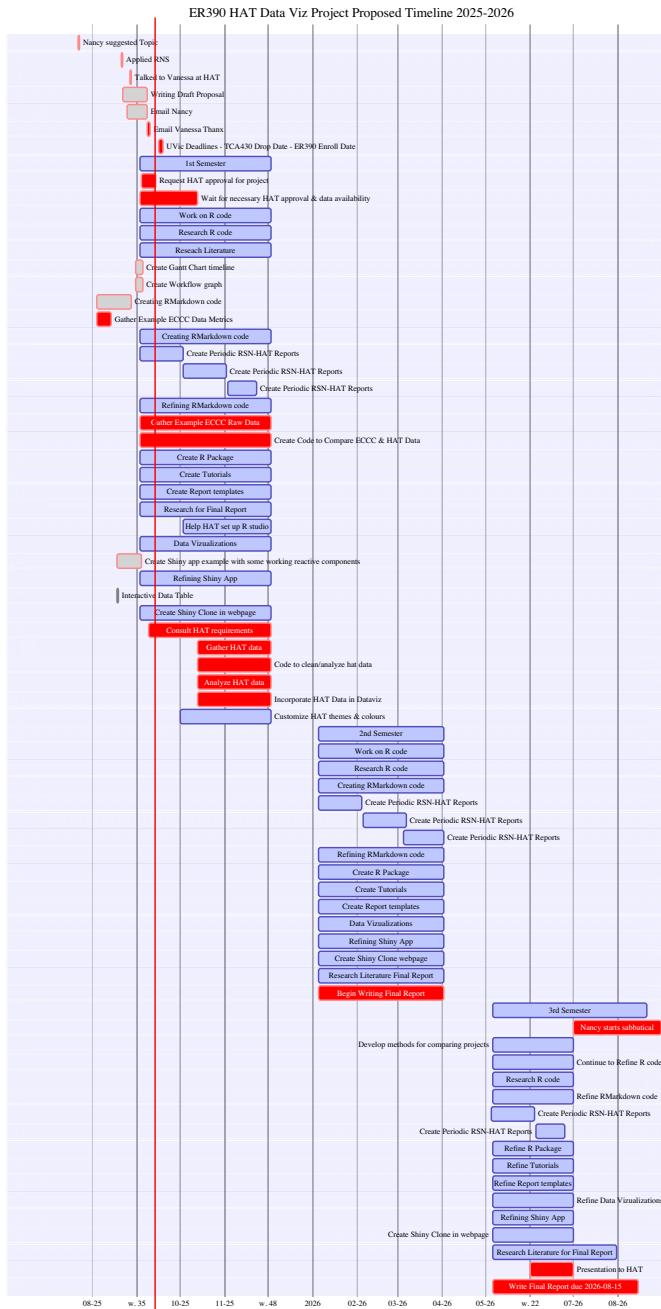


Figure 2: Gantt Chart for Proposed ER390 Timeline, Sep 2025-Jul 2026

---

## Data Cleaning and Transformation Example Results

### Data download and formatting

- These example visualizations use GOE monitoring data from report by Malloff & Shackelford (2024)
- Data (in excel or csv format) is uploaded
- Data is formatted, saved, and transformed to be useful for creating visualizations
- The code chunks have been hidden for the HTML report, highlighting only the resulting plots, though code becomes visible in PDF document. The R code has been documented using hidden comments.

### Load Site Data

Click open to see R code for Loading Data

```
# Load data and create 2 data frames
# Load Site Data
goe_eccc <- read.csv("data/SiteData-TEST-GOE-Monitor.csv", header = TRUE, sep = ",")
# Load Site Details
Sites_eccc <- read.csv("data/SiteDetails-TEST-GOE-Monitor.csv", header = TRUE, sep = ",")
```

### Data Cleaning and Transformation

Click open to see R code for Data Cleaning and Transformation

```
# create new column with proportions changed to percentage as a whole number
Percentage_ns <- goe_eccc$Proportion_of_native_species * 100

# add-new column for Percentages
goe_eccc <- cbind(goe_eccc, Percentage_ns)

# arrange row order of site names for binding 2 dataframes
Sites_eccc_1 <- Sites_eccc %>% arrange(Sites_eccc$Site)
goe_eccc_1 <- goe_eccc %>% arrange(goe_eccc$Site)

## make sure both data file columns have same data type str()
## make sure they have the same number of rows e.g. length(unique(goe_eccc_1$Site))

# bind 2 dataframes together by columns
goe_eccc_sites <- cbind(Sites_eccc_1, goe_eccc_1)

# Remove duplicate columns resulting from binding 2 data frames
goe_eccc_sites_1 <- goe_eccc_sites[c(-9,-10)]

# Rename columns to remove dot-format
names(goe_eccc_sites_1)[names(goe_eccc_sites_1) == "Land.Manager"] <- "Land_Manager"
names(goe_eccc_sites_1)[names(goe_eccc_sites_1) == "Main.Restoration.Type"] <- "Main_Restoration_Type"
names(goe_eccc_sites_1)[names(goe_eccc_sites_1) == "Restoration.Intensity"] <- "Restoration_Intensity"
names(goe_eccc_sites_1)[names(goe_eccc_sites_1) == "Area.ha"] <- "Area_ha"

# write csv file with combined data
## creates new data file used for creating visualizations
write.csv(goe_eccc_sites_1, "data/goe_eccc_all_site_data.csv", row.names = FALSE)
```

# Data Visualization Example Results

## Data Table

- Data from Malloff & Shackelford (2024)

### Full Data Table

- The full table is too wide for use in pdf because current variable names are too long, and unlike HTML, PDFs do not allow sideways scrolling.

Subregion Site Land\_Manager Main\_Restoration\_Type Restoration\_Intensity Area\_ha Lat Lng Proportion\_of\_native\_species Cultural\_species\_richness Exotic\_species Trampling Herbivory Composite\_Index Year Percentage\_ns

### Table Head of Site Information

Subregion	Site	Land.Manager	Main.Restoration.Type	Restoration.Intensity	Area.ha	Lat	Lng
Gulf Islands	Anniversary Island	GINPR	herbivore reduction	high	4.39	48.82292	-123.1823
Gulf Islands	AVNR	Saltspring conservancy	herbivore reduction	minimal	20.54	48.80351	-123.4425
Saanich Peninsula	Bear Hill Park	CRD	invasive removal	low	3.80	48.54639	-123.4078
Gulf Islands	Brackman Island	GINPR	invasive removal	high	4.41	48.71897	-123.3864
Saanich Peninsula	Camas Hill	HAT	invasive removal	high	10.10	48.40173	-123.5975
Gulf Islands	Crows Nest	Trinity Western University	herbivore reduction	high	15.34	48.78237	-123.4612

### Table Head of Site Data

Subregion	Site	Proportion_of_native_species	Cultural_species_richness	Exotic_species	Trampling	Herbivory	Composite_
Gulf Islands	Anniversary Island	0.94	2.71	0.17	5.00	1.14	
Gulf Islands	AVNR	0.55	1.30	57.19	4.70	36.50	
Saanich Peninsula	Bear Hill Park	0.62	1.86	24.46	7.57	9.50	
Gulf Islands	Brackman Island	0.77	2.80	11.17	5.40	0.00	
Saanich Peninsula	Camas Hill	0.61	2.43	27.19	7.71	4.71	
Gulf Islands	Crows Nest	0.66	1.07	31.51	1.64	26.71	

Figures 3,4,5: Example Data Table Formats, with Table Head showing only first 6 rows.

## Map of Site Locations

### Leaflet Map with Circle Markers

- The Leaflet Map is created using R code in RMarkdown.
- Data is summarized in map marker popups with summary of site data (Malloff & Shackelford, 2024).

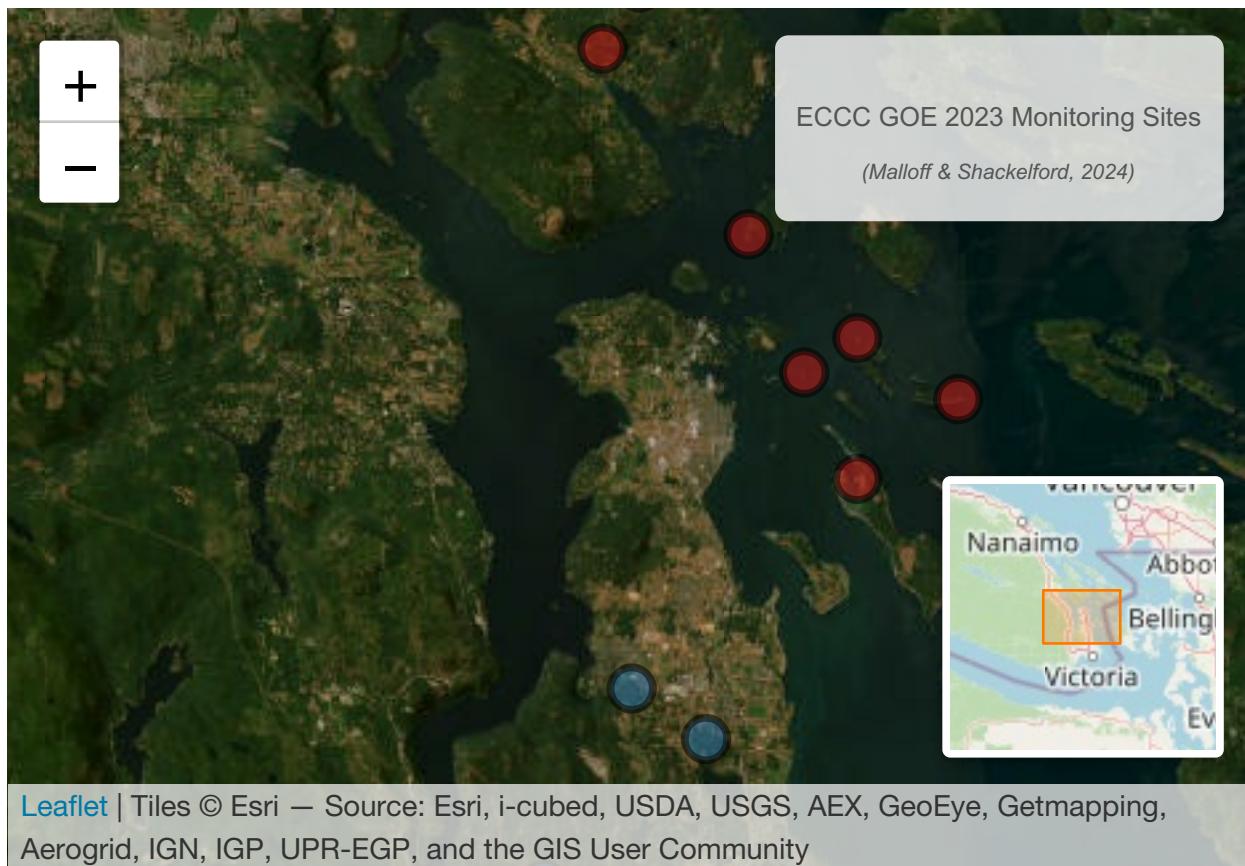


Figure 6: Interactive Leaflet Map with Example Sites Data (Malloff & Shackelford, 2024).

## Subregion Plots

- These plots compare monitoring variable measurement results between two Subregions.
- Data from Malloff & Shackelford (2024).

### Violin/Bar Plot: Compare Subregions Exotic and Native Species

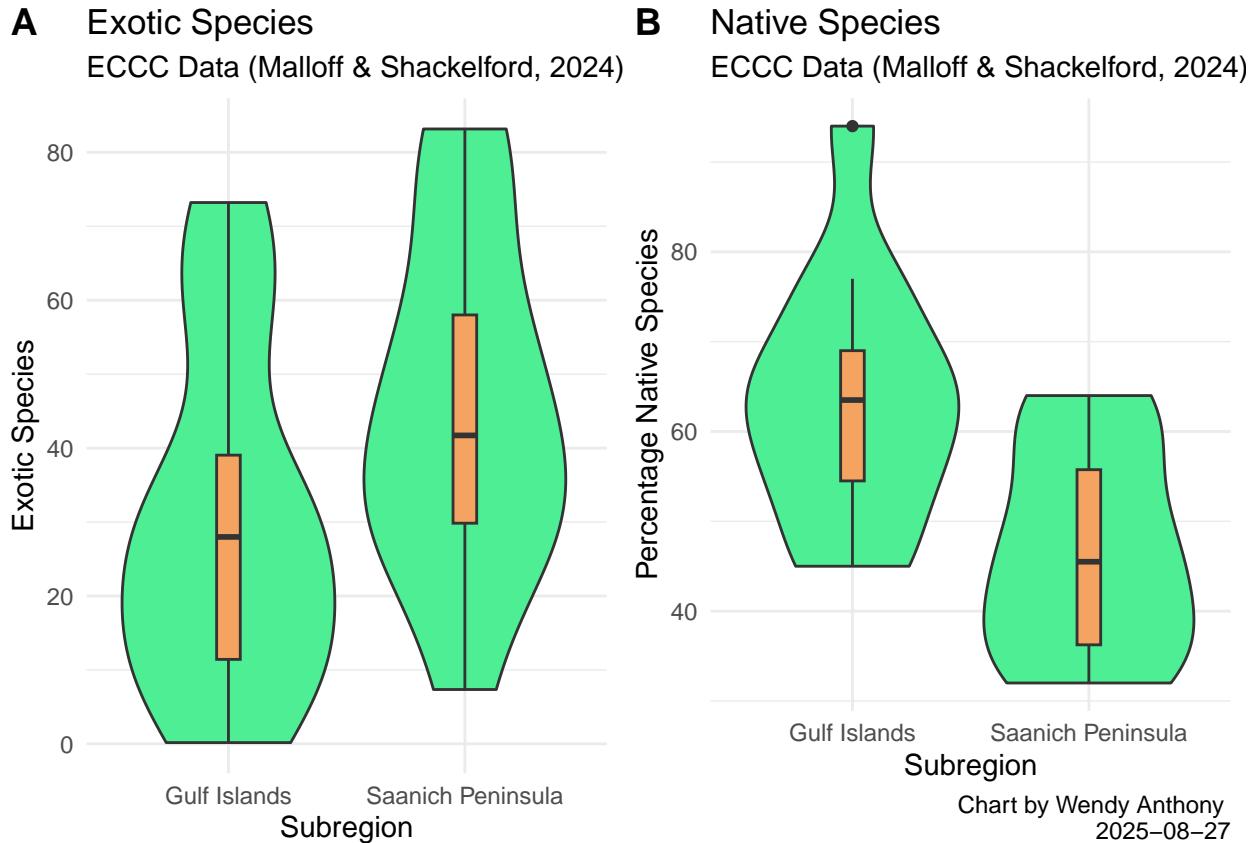


Figure 7: Violin / bar plot compares variables of 2 Subregions

## Site Plots

### Point Plot: Compare Individual Sites by Restoration Intensity

- Sites are ordered by the level of restoration intensity, ranging from high to minimal.
- This plotly-style chart becomes interactive in html version by clicking data points for view a tool-tip of data results.
- Data from Malloff & Shackelford (2024).

### Plot Site Restoration Intensity (High to Low)

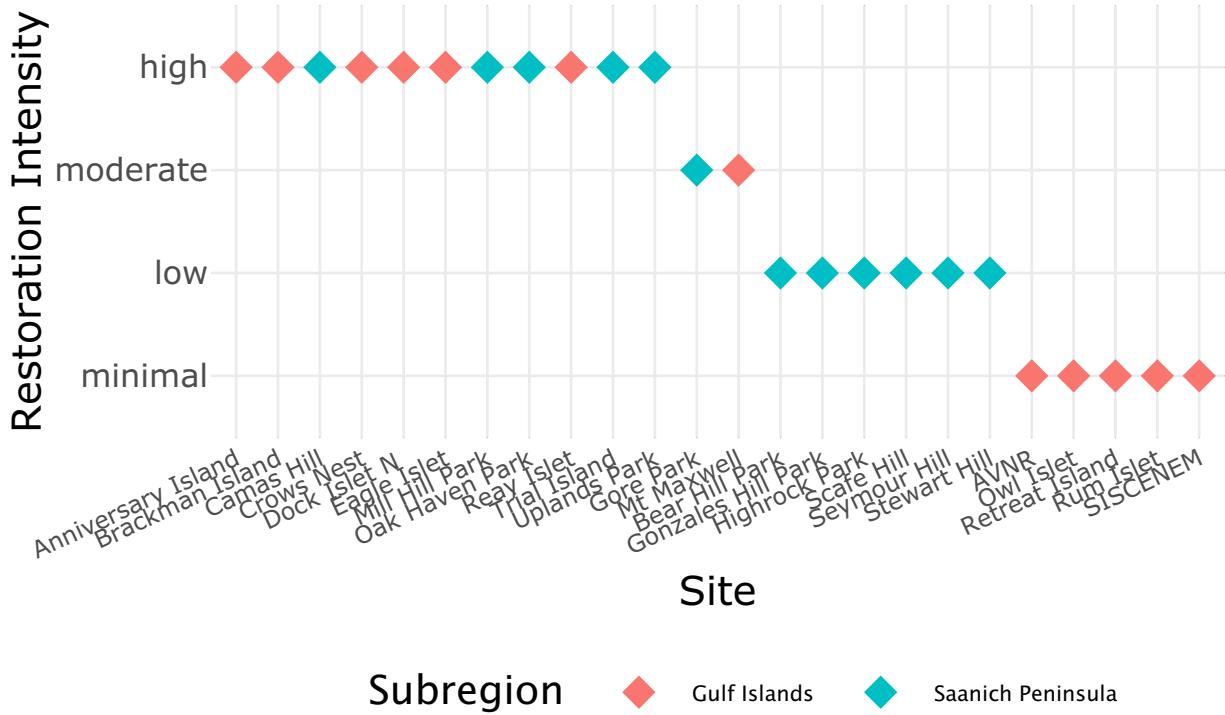


Figure 8: Point Plot of Restoration Intensity of Example Sites, coloured by Subregion.